

# GCBN Gene Annotation, Analyses and Data HMGU Provision Services for Cereal Genomes

Heidrun Gundlach, Thomas Lux, Daniel Lang, Georg Haberer, Manuel Spannagl, Klaus FX Mayer  
PGSB, Plant Genome and Sytems Biology, Helmholtz Center Munich

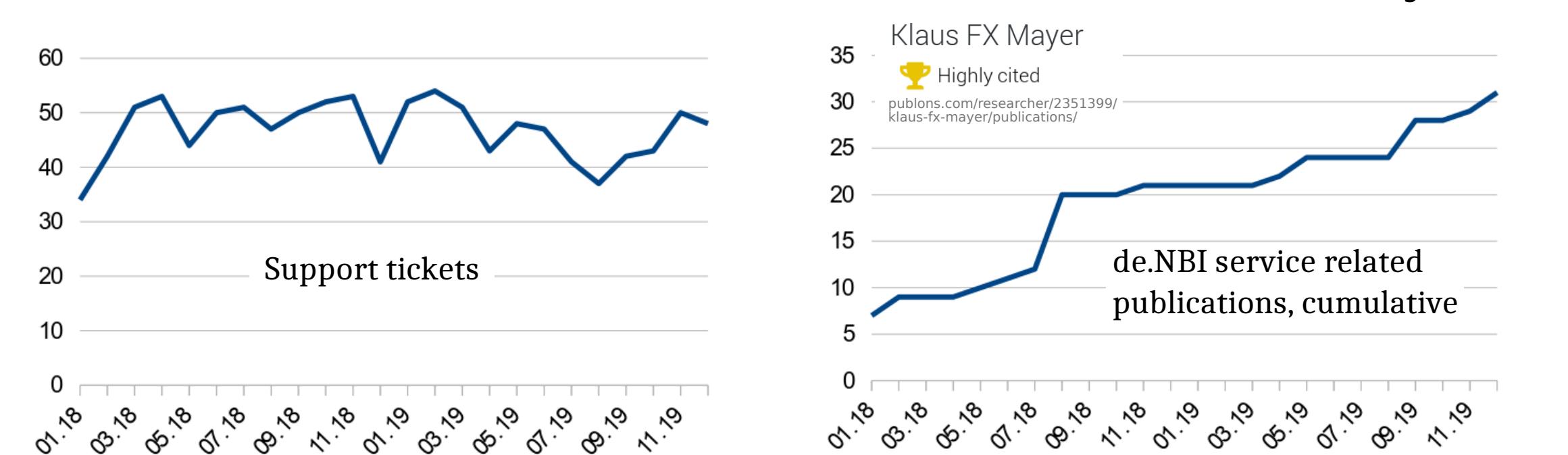
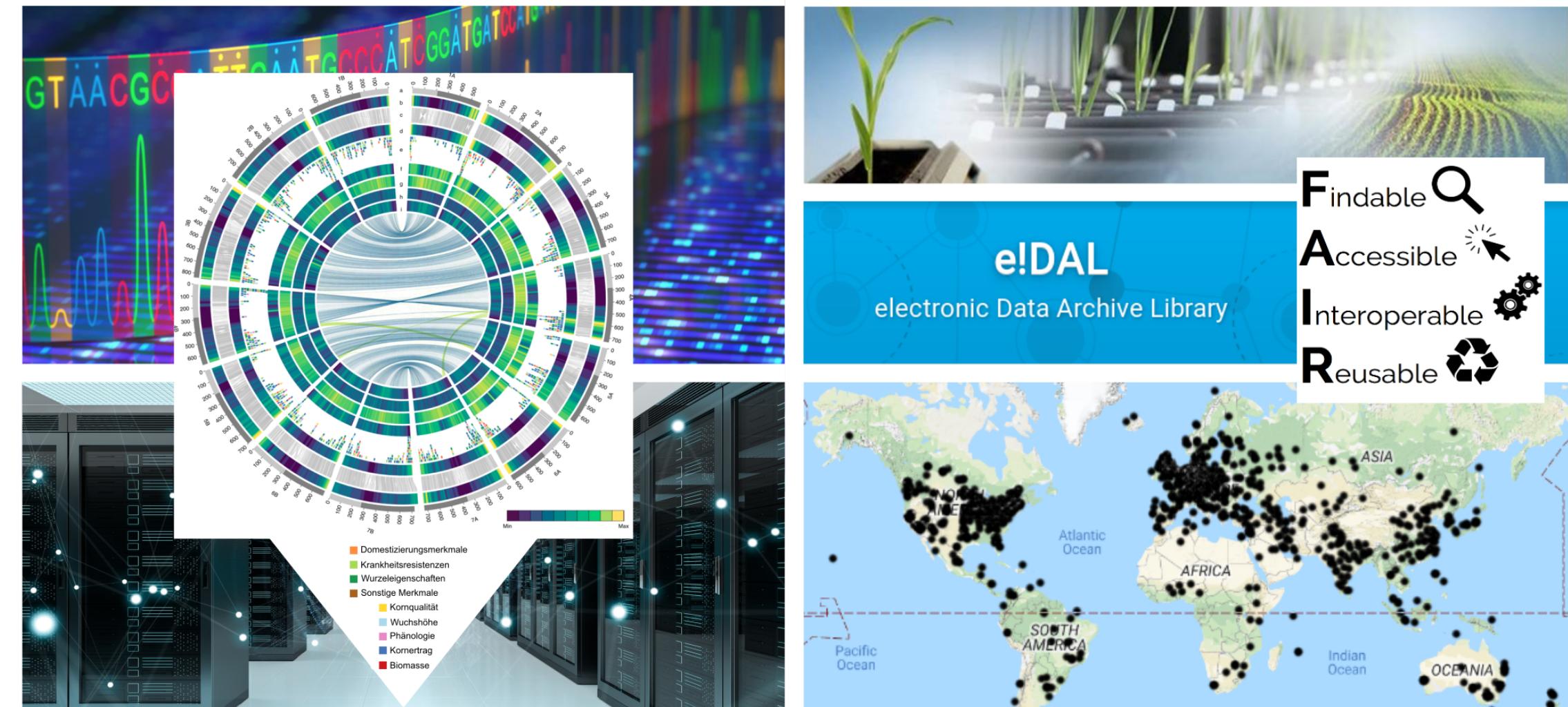
Fkz031A536B

## Short description of the project

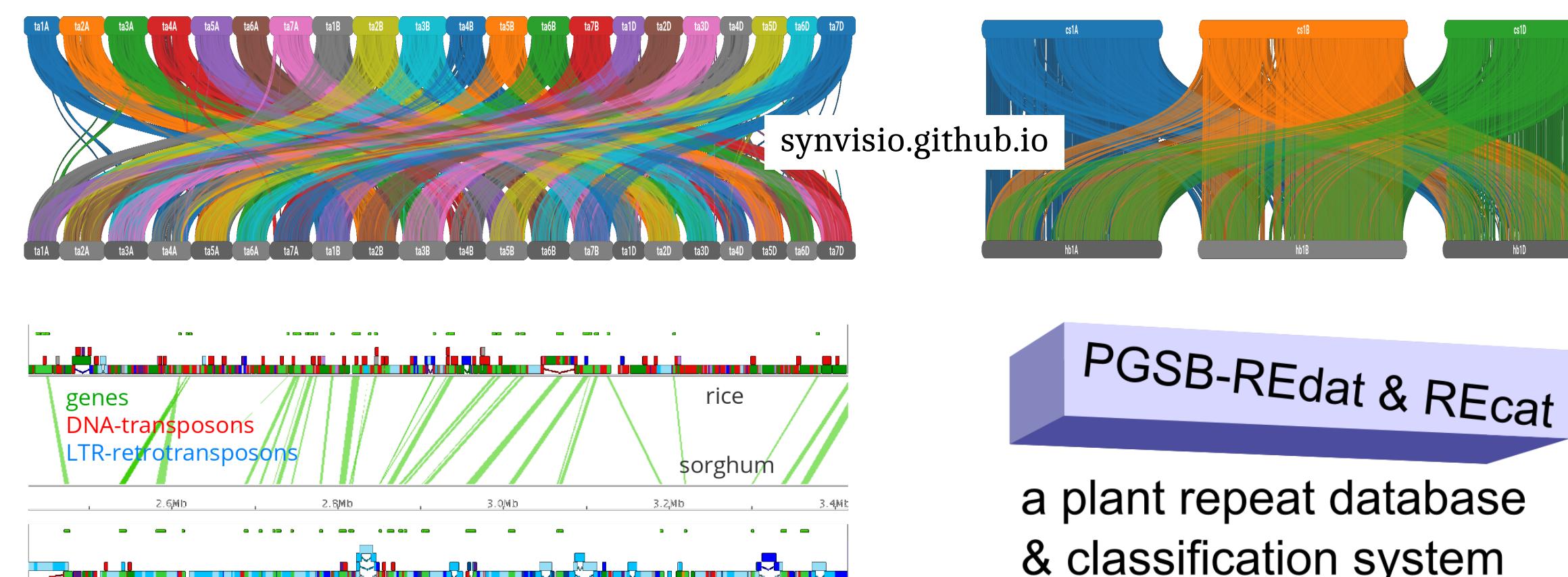
Our tasks within the service center GCBN (German Crop BioGreen-formatics Network) are the development and improvement of workflows for gene annotation in plants, with a focus on the large and complex cereal genomes. We provide comprehensive gene annotations and subsequent analyses, often on the basis of cooperations within international genome projects. In addition we do consulting as service to the plant community. This service covers for instance consulting on best practices and approaches for plant genome analyses, the search for suitable datasets and their interpretation or the extraction of customized datasets.

## de.NBI services

- timely delivery of annotations for large crop genomes (is not trivial: such complex and transposon crowded genomes require sufficient computing resources & optimized pipelines)
- FAIR data deposition
- consulting service on best practices for plant genomes



Upcoming resources: interactive and scalable pan-genome visualisation and plant transposon database



## General information on the project

- 1 scientist FTE financed by de.NBI
- further staff involved with de.NBI tasks add up to 1 FTE:  
0.1 PI (K.Mayer), 0.5 pipeline implementation (T.Lux, Postdoc),  
0.2 FTE system administration, 0.2 smaller service tasks

## Progress report

[github.com/PGSB-HMGU/plant.annot](https://github.com/PGSB-HMGU/plant.annot)

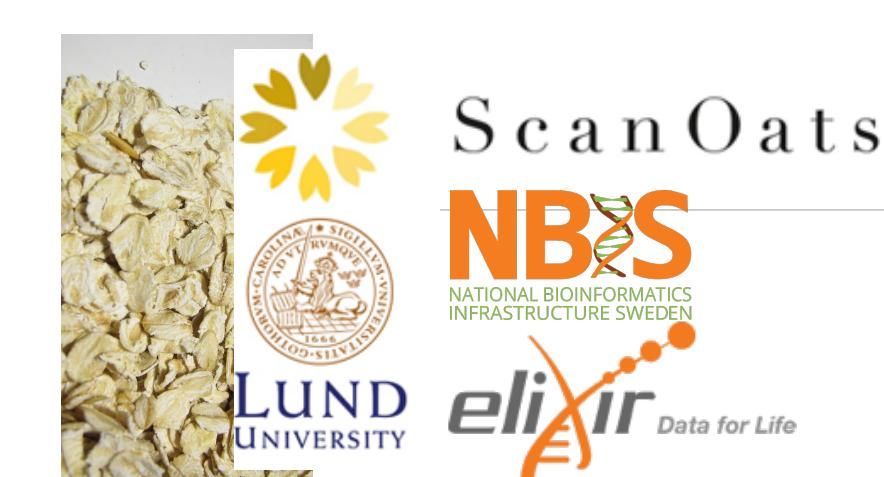
### Annotations for several (Pan-)genome projects (270 Gb sequence)



OPENWILDWHEAT



Rye genome project



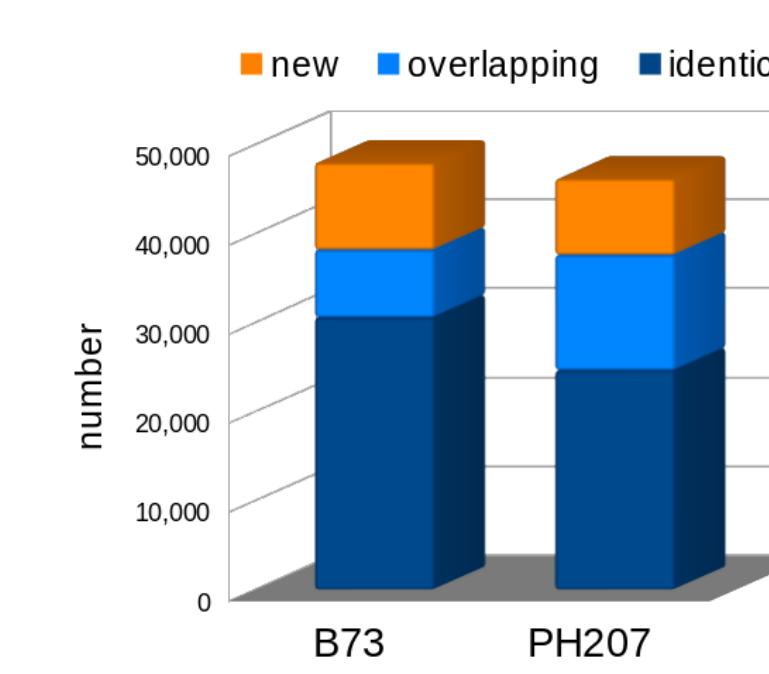
MaizeGDB 4 European lines

### Gene consolidation across pan-genomes to identify missing genes

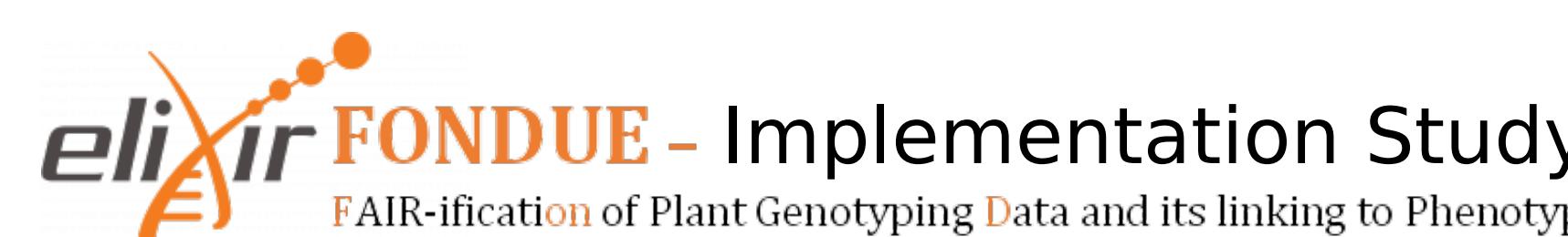
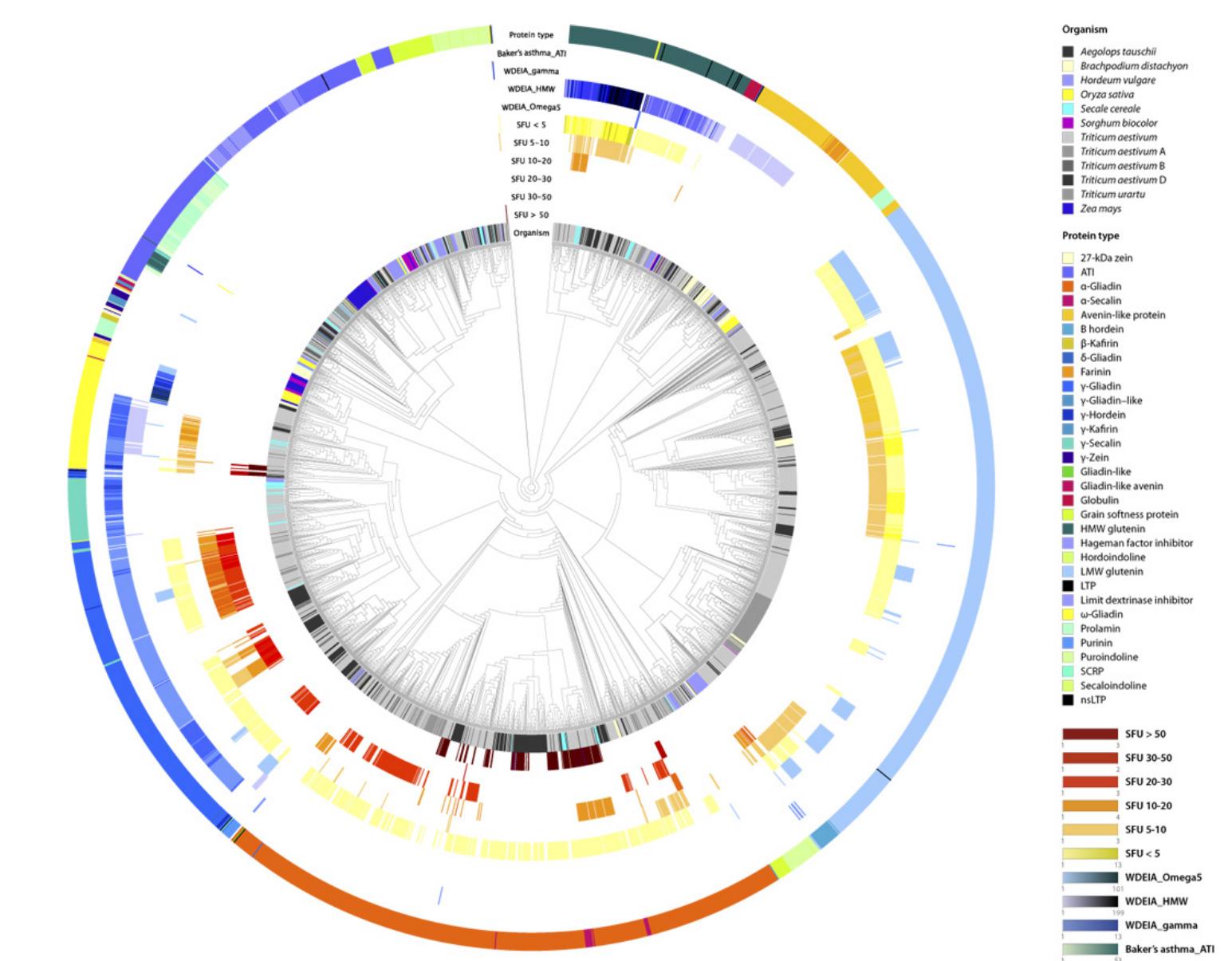
Thomas Lux: "A consolidated gene annotation for the 10+ wheat genome project"



Increase of gene numbers after consolidation (example maize)



Wheat allergene atlas as example for studies based on gene annotations (Juhasz et al. 2018)



## de.NBI training and education (2019)

- individual training of PHD students to use the HMGU gene-pipeline in their project (e.g. Oat)
- use of de.NBI cloud virtual servers as common platform in a bioinformatic course for university students

## Publications (de.NBI related, since 2018 SAB meeting)

- HMGU (gene) annotation      ● method description
- M. Timothy Rabanus-Wallace, Bernd Hackauf, Martin Mascher, Thomas Lux, Thomas Wicker, Heidrun Gundlach, Mariana Baez, Andreas Houben, Klaus FX. Mayer, ... Uwe Scholz, Alan H. Schulman, Dörthe Siekmann, Stefan Stojalaowski, Vijay Tiwari, Manuel Spannagl, Nils Stein. Chromosome-scale genome assembly provides insights into rye biology, evolution, and agronomic potential. *BioRxiv*. 2019 Dec 12. doi: 10.1101/2019.12.11.869693
- Stam R, Nosenko T, Hörger AC, Stephan W, Seidel M, Kuhn JMM, Haberer G, Tellier A. The *de Novo* Reference Genome and Transcriptome Assemblies of the Wild Tomato Species *Solanum lycopersicum* Highlights Birth and Death of NLR Genes Between Tomato Species. *G3*. 2019 Dec doi: 10.1534/g3.119.400529.
- Gardiner LJ, Joynson R, Omony J, Rusholme-Pilcher R, Olohan L, Lang D, Bat C, Hawkesford M, Salt D, Spannagl M, Mayer KFX, Kenny J, Bevan M, Hall N, Hall A. Hidden variation in polyploid wheat drives local adaptation. *Genome Res.* 2018 Sep;28(9):1319-1332. doi: 10.1101/gr.233551.117.
- G. Haberer, E. Bauer, N. Kamal, H. Gundlach, I. Fischer, M.A. Seidel, M. Spannagl, C. Marcon, A. Ruban, C. Urbany, A. Nemri, F. Hochholdinger, M. Ouzunova, A. Houben, C.-C. Schön, K.F.X. Mayer. European maize genomes unveil pan-genomic dynamics of repeats and genes. *BioRxiv*. 2019 Sep 13. doi: 10.1101/766444
- Yoshida S, Kim S, Wafuda EK, Tanskanen J, ... Cui S, Der JP, Gundlach H, Jao Y, Horváth C, Ishida JK, ... Schulman AH, Timko MP, dePamphilis CW, Choi D, Shirasu K. Genome Sequence of *Striga asiatica* Provides Insight into the Evolution of Plant Parasitism. *Curr Biol*. 2019 Sep 23;29(18) doi: 10.1016/j.cub.2019.07.
- Maccaferri M, Harris NS, Twardziok SO, Pasam RK, Gundlach H, Spannagl M, Ormanbekova D, Lux T, Prade VM, Milner SG, Himmelbach A, Mascher ... Stein N, Hale I, Distefano A, Hayden MJ, Tuberrosa R, Walkowiak S, Mayer KFX, Cerotti A, Pozniak CJ, Cattivelli L. Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat Genet*. 2019 May;51(5):885-895. doi: 10.1038/s41588-019-0381-3.
- Pont C, Leroy T, Seidel M, Tondelli A, Duchemin W, Armisén D, Lang D, Bustos-Korts D, Goué N, Balfourier F, Mohr-Láng M, Lage J, Kilian B, Özkan H, Waite D, Dyer S, Letellier T, Alaux M. Wheat and Barley Legacy for Breeding Improvement (WHEALIB) consortium, Russell J, Keller B, van Eeuwijk F, Spannagl M, Mayer KFX, Waugh R, Stein N, Cativelli L, Haberer G, Charmet G, Salse J. Tracing the ancestry of modern bread wheats. *Nat Genet*. 2019 May;51(5):905-911.
- Omony J, Nussbaumer T, Gutztav R. DNA methylation analysis in plants: review of computational tools and future perspectives. *Brief Bioinform*. 2019 Apr 9.
- Monat C, Padmarasu S, Lux T, Wicker T, Gundlach H, Himmelbach A, Ens J, Li C, Muehlbauer GJ, Schulman AH, Waugh R, Braumann I, Pozniak C, Scholz U, Mayer KFX, Spannagl M, Stein N, Mascher M. TRITEX: chromosome-scale sequence assembly of Triticeae genomes with open-source tools. *Genome Biol*. 2019 Dec 10;20(1):290.
- Bolger AM, Poorter H, Dumschat K, Bolger ME, Arend D, Osorio S, Gundlach H, Mayer KFX, Lange M, Scholz U, Usadel B. Computational aspects underlying genome to phenotype analysis in plants. *Plant J*. 2019 Jan;97(1):182-198. doi: 10.1111/pj.14179.
- Keilwagen J, Lehner H, Berner T, Beier S, Scholz U, Himmelbach A, Stein N, Badaeva ED, Lang D, Kilian B, Hackauf B, Perovic D. Detecting Large Chromosomal Modifications Using Short Read Data From Genotyping-by-Sequencing. *Front Plant Sci*. 2019 Sep 24;10:1133. doi: 10.3389/fpls.2019.01133.
- Messerer, M.; Lang, D.; & Mayer, K.F.X. Analysis of stress resistance using next generation techniques. *Agronomy* 8:130 (2018)
- Thind AK, Wicker T, Müller T, Ackermann PM, Steuernagel B, Wulff BBH, Spannagl M, Twardziok SO, Felder M, Lux T, Mayer KFX; International Wheat Genome Sequencing Consortium, Keller B, Kräutler S. Chromosome-scale comparative sequence analysis unravels molecular mechanisms of genome dynamics between two wheat cultivars. *Genome Biol*. 2018 Aug 17;19(1):104. doi: 10.1186/s13059-018-1477-2.
- Juhasz A, Belova T, Florides CG, Maulis C, Fischer I, Gell G, Birinyi Z, Ong J, Keeble-Gagnière G, Maharajan A, Ma W, Gibson P, Jia J, Lang D, Mayer KFX, Spannagl M; International Wheat Genome Sequencing Consortium, Keller B, Kräutler S. Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat. *Sci Adv*. 2018 Aug 17;4(8):ear8602. doi: 10.1126/sciadv.aar8602.
- Keilwagen J, Hartung F, Paulini M, Twardziok SO, Grau J. Combining RNA-seq data and homology-based gene prediction for plants, animals and fungi. *BMC Bioinformatics*. 2018 May 30;19(1):189. doi: 10.1186/s12859-018-2203-5.